Forum for Linguistic Studies | Volume 07 | Issue 02 | February 2025

BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Personalizing the Making of Technical Word Lists for Science and Technology Students: A Pedagogic Action Research Study

Ruchi Yadav [1*] **,** ***Shakul Tewari*** [2] **,** ***Vijayakumar Chintalapalli*** [1]

[1] *Department of Humanities and Social Sciences, Birla Institute of Technology and Science (Pilani), Rajasthan 333031, India*

[2] *Department of Humanities and Social Sciences, Birla Institute of Technology and Science (Hyderabad), Telangana 5000078, India*

ABSTRACT

The challenge of meeting the diverse and specific vocabulary needs of students from different disciplines has been widely discussed in academic research. This paper explores how undergraduate students in science and engineering personalized the process of creating word lists for their academic needs. By using two corpus analysis tools—wordlist and concordance—available through the SketchEngine interface, the students were able to analyze an ad hoc corpus of texts relevant to their fields. Additionally, they assessed the technicality of words using a technicalness scale specifically developed for the purpose of this study. The data for the study consisted of the word lists that the students submitted, along with their notes, and transcriptions of individual student presentations on these word lists. The findings revealed that the students categorized words based on their technicality, considering both the specific context within their academic corpus and the broader meanings that these words could hold outside of their disciplines. In particular, students distinguished words based on their level of specialization and how these terms were understood in other fields. These results underscore the importance of a discipline-specific approach when teaching vocabulary in science and engineering programs. In conclusion, we discuss potential methods for teaching vocabulary to science and technology students, emphasizing the importance of personalized learning and context-based vocabulary development to support academic success.

Keywords: EAP; Concordance; Peer-Discussion; Technicalness Scale; Word Lists

*CORRESPONDING AUTHOR:

Ruchi Yadav, Department of Humanities and Social Science, Birla Institute of Technology and Science (Pilani), Rajasthan 333031, India; Email: p20210437@pilani.bits-pilani.ac.in

ARTICLE INFO

Received: 23 December 2024 | Revised: 6 February 2025 | Accepted: 8 February 2025 | Published Online: 26 February 2025

DOI: https://doi.org/10.30564/fls.v7i2.8194

CITATION

Yadav, R., Tewari, S., Vijayakumar, C., 2025. Personalizing the Making of Technical Word Lists for Science and Technology Students: A Pedagogic Action Research Study. Forum for Linguistic Studies. 7(2): 1151–1015. DOI: https://doi.org/10.30564/fls.v7i2.8194

1. Introduction

Research has shown that understanding and using academic and technical vocabulary determines the quality of engagement with the disciplinary discourses in higher education (HE)[1–3]. Although there is no consensus on the exact number of words a student is required to learn to be successful in HE, recent estimates highlight the view that a large number of content words in academic prose are polysemous and context-dependent and their frequency 'cut across levels' and often expressing specialized meanings that are barely used in everyday communication[4]. Research shows that these context-sensitive and specialized meanings cause comprehension issues for novice students who have to engage with the disciplinary discourses regularly[5].

Keeping this in view several academic word lists have been produced to facilitate student learning. For example, a focus on general academic vocabulary resulted in the publication of lists such as the Academic Word List (AWL) and Academic Vocabulary List (AVL)[1, 6], and a focus on discipline-specific academic vocabulary resulted in the production of discipline-specific vocabulary lists[7, 8]. In terms of serving the academic vocabulary needs of specific disciplines, these word lists were highly useful[9, 10]. However, they were either too general covering a wide range of disciplines or specific to any one discipline although they were based on large representative sample of academic texts. That is, as Hyland and Tse[2] have observed none of these lists constituted 'a single restricted lexical repertoire'that an individual student could rely on. Given the fact that many undergraduate students, in this case science and engineering students, pursue multi- and cross-disciplinary courses, they need tailor-made word lists that serve their specific academic vocabulary needs. Therefore, we have taken up this study to help the learners to identify the lexical resources they need for their academic purposes.

Furthermore, teaching a comprehensive list of vocabulary items in technical universities is constrained by the curricular requirements. At best, we can teach the students a set of strategies and provide them with the resources to practice their vocabulary. Keeping this systemic constraint in mind, we have devised a plan of action around corpus-based pedagogy[11]. Students were introduced to corpus-analysis tools along with a technicalness scale developed for the purpose of analysing the words. This paper discusses how a group of undergraduate engineering students majoring in a variety of science and engineering disciplines compiled their own discipline/topic-specific corpora, which contained texts/readings suggested for study by the professors whose courses the students have opted for, and analysed the vocabulary using the technicalness scale. The following research questions were addressed.

2. Research Questions

(1) How pervasive are the AWL words in the corpora put together by the students?
(2) How did the students classify the words into semi-technical and technical?

3. Review of Literature

Research on vocabulary in English for Academic Purposes (EAP) has primarily focused on two types of word lists: (1) general academic word lists and (2) discipline-specific academic word lists. These lists aim to facilitate student learning by providing lexical resources tailored to academic contexts. The Academic Word List (AWL)[6], the Academic Vocabulary List (AVL)[1], and the Academic Collocation List[12] were developed to support general academic vocabulary acquisition. In contrast, domain-specific lists, such as the Plumbing Word List[7], the Chemistry Academic Word List[13], and the Nursing Academic Word List[8], target specialized vocabulary needs. These lists have been widely used in EAP contexts, as they offer a structured approach to teaching lexical items that are essential for academic success. However, research has raised concerns about the applicability and limitations of these lists. Hyland & Tse[2] argue that no single lexical repertoire can fully address students' academic vocabulary needs, as word lists are often either too general (covering multiple disciplines) or too restrictive (focusing on a single field). This issue is particularly relevant for science and engineering students, who frequently engage in cross-disciplinary work and require a more flexible, context-sensitive approach to vocabulary learning[9, 10].

3.1. Context-Sensitive and Corpus-Based Approaches

To address these concerns, researchers have explored corpus-based approaches to vocabulary teaching. Studies have shown that context-sensitive word lists, developed using corpus analysis tools, offer a more tailored approach to EAP vocabulary instruction. For example, register variation studies have demonstrated significant differences in word usage across academic genres, such as lectures vs. research articles or conference proceedings vs. textbooks [14–17].

This shift toward discipline-specific corpora has led to a growing emphasis on lexical profiling in STEM disciplines [18, 19]. Rather than treating vocabulary as a static list, researchers advocate for a dynamic, student-generated approach, where learners compile their own discipline-specific corpora to identify key academic and technical terms relevant to their studies [20, 21]. This approach aligns with data-driven learning (DDL) methodologies, which encourage students to explore word usage patterns through corpus tools such as SketchEngine and AntConc [22, 23].

3.2. Defining "Technical" and "Sub-Technical" Vocabulary

A critical challenge in vocabulary research is defining what constitutes "technical" vs. "academic" vocabulary. Traditionally, vocabulary classification has relied on frequency-based methods, where words are categorized into bands (e.g., 1k, 2k, and 3k levels) based on their occurrence in academic texts [10, 24, 25].

However, recent studies argue that frequency alone is insufficient for determining technical vocabulary [26]. Words may exhibit context-dependent technicality, where common words take on specialized meanings within specific disciplines [27]. For example, in aviation, words such as *request*, *feet*, and *grounding* carry highly specialized meanings distinct from their everyday usage [26]. Similarly, in chemistry, words like *spectrum*, *crystal*, and *compound* acquire discipline-specific meanings that differ from general academic usage [13].

To account for variations in word technicality, researchers have proposed context-sensitive classification models. Chung & Nation [28, 29] developed a sense-based rating scale, categorizing words based on their occurrence within and beyond a discipline. Their model identifies four levels of technicality, distinguishing between general academic words, sub-technical words, and highly technical terms.

More recently, Ha & Hyland [27] refined this approach by incorporating polysemy, collocation patterns, and register variation into their classification system. Their Technicality Analysis Model (TAM) assigns words to five categories based on their semantic distance from general usage. This model demonstrates that certain words project discipline-specific meanings while maintaining general academic use in other contexts. In this study, we build on these frameworks by developing a Technicalness Scale (**Figure 1**) that categorizes vocabulary into three primary groups: General Academic, Sub-Technical, and Technical. This model provides students with a structured heuristic for analyzing word technicality across disciplines.

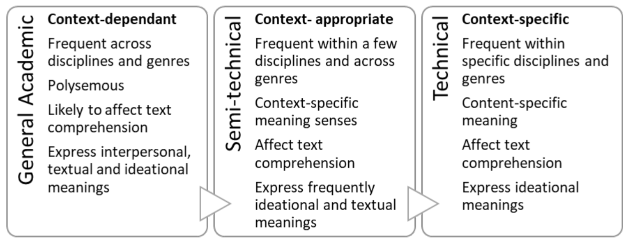


Figure 1. Model of technicalness scale for identifying the general academic, sub-technical and technical vocabulary.

As shown in **Figure 1**, the model classifies words based on their context-dependence, frequency of occurrence in specific disciplines, and their role in conveying ideational, textual, or interpersonal meanings. Unlike previous models, our scale integrates corpus-based student engagement to refine word categorization through peer discussion. This approach allows students to actively determine the technicality of vocabulary items rather than relying solely on pre-existing word lists. In the following sections, we discuss how this scale was implemented in student-led corpus analysis and peer discussions to enhance vocabulary classification.

3.3. Peer Discussion (PD) and Student-Generated Word Lists

While most EAP vocabulary studies focus on pre-compiled word lists, recent research has highlighted the benefits of student-led corpus analysis. Learner autonomy plays a crucial role in vocabulary acquisition, as students actively engage with texts to determine the technicality of words relevant to their disciplines [4, 30].

Studies have shown that peer discussion enhances the accuracy of student-generated word lists. When students collaborate to categorize words as technical or sub-technical, they develop a critical awareness of lexical variation and discipline-specific usage [31, 32]. Peer engagement has been found to:

- Lower affective barriers to learning [33].
- Promote metacognitive awareness of vocabulary selection [34].
- Encourage disciplinary dialogue, where students justify their word choices through examples and counterexamples [35].

However, few studies have systematically documented the process of student-led corpus building in EAP. While research supports the pedagogical benefits of personalized vocabulary learning, little is known about how students negotiate technicality classification in real-world contexts. This study aims to fill this gap by exploring how engineering students categorize technical and sub-technical vocabulary through peer discussion and corpus analysis.

4. Methodology

This study employs a pedagogic action research approach to explore how undergraduate STEM students categorize technical and sub-technical vocabulary through peer discussion and corpus analysis. The research integrates student-led corpus building, data-driven learning (DDL) methodologies, and a modified technicalness scale to enhance discipline-specific vocabulary learning. This section details the participants, data collection methods, corpus compilation, word categorization process, and data analysis procedures.

4.1. Participants and Context

This study was conducted with 39 undergraduate students enrolled in the English Skills for Academics (HSS F224) course at Birla Institute of Technology and Science – Pilani. The participants represented diverse STEM disciplines, including Computer Science, Biochemistry, Physics, Mechanical Engineering, Pharmacology, and Applied Economics. All students were non-native English speakers with at least eight years of English education. The course followed a blended learning format, with 42 instructional hours, including five hours dedicated to corpus-based vocabulary analysis using SketchEngine.

The sample size of 39 students was determined by the availability of students enrolled in the course. While larger samples are often preferred for generalizability, action research in educational settings frequently operates with naturally occurring classroom groups [36, 37]. Research in corpus-driven EAP pedagogy often utilizes small participant groups due to the intensive nature of data collection and analysis [38, 39]. Lee and Swales [40] also highlight that rich corpus data compensates for small sample sizes, as it provides deep insights into disciplinary language use.

Since vocabulary learning needs vary across disciplines, each student was required to compile a discipline-specific corpus and generate word lists relevant to their field of study. To ensure peer learning and cross-disciplinary engagement, students worked in pairs or small groups, engaging in structured peer discussions to classify vocabulary items as general academic, sub-technical, or technical.

4.2. Corpus Compilation and Word List Generation

Each student compiled a Corpus of Expert Writing (CEW) using 20–30 e-texts from research articles, textbook chapters, review papers, and conference proceedings relevant to their discipline. The compiled texts were processed using SketchEngine, a corpus analysis tool that allowed students to:

(1) Extract word lists based on frequency and range.
(2) Analyze concordance lines to determine word usage in different contexts.
(3) Identify collocations and phraseological patterns of key vocabulary items.

After generating raw word lists, students manually cleaned the data, removing proper nouns, email IDs, symbols, and non-academic words. The resulting 29-word lists (after removing data from students who withdrew) formed the basis of the study's vocabulary classification.

The final Corpus of Expert Writing (CEW) consisted of 29 discipline-specific sub-corpora, comprising 8,494,046 tokens and 372 headwords from the Academic Word List

(AWL)[6].

Table 1 presents the distribution of AWL words across sub-corpora, highlighting disciplinary variations in vocabulary usage.

Table 1. CEW corpus compiled by the students.

Corpus of Expert Writing	Count
Number of sub-domains/disciplines	29
Number of tokens	8494046
Number of AWL words	372

4.3. Vocabulary Categorization Using the Technicalness Scale

As per the task requirements, vocabulary modules had four major teaching objectives. The first one was to generate a word list; the second one was to publish academic and discipline-specific word lists (sub-technical & technical). The third objective was to engage in a dialogue with their peers to finalize the technicalness of the identified academic words. It required the students to discuss and justify to their peers why certain words were classified under academic or semi-technical or technical. And the peers' academic backgrounds varied considerably from each other. Finally, the students were asked to make individual presentations on the word lists produced. The process of producing the academic and technical vocabulary is given in **Figure 2**.

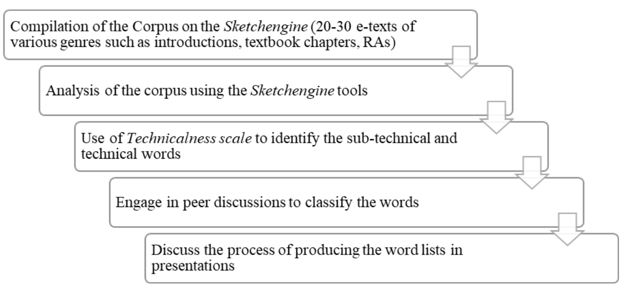


Figure 2. Method of making the word lists.

Students categorized words into three primary groups using the Technicalness Scale (**Figure 1**, Section 3). The classification process involved:

1. Identifying General Academic Words – High-frequency words appearing across multiple disciplines (e.g., *method, approach, significant*).
2. Determining Sub-Technical Words – Context-dependent words frequently used within a few disciplines but retaining some general meanings (e.g., *crystal, spectrum, grounding* in Chemistry and Engineering).
3. Classifying Technical Words – Highly specialized vocabulary restricted to specific disciplines (e.g., *nanosheet, electrolyte, homomorphic encryption*).

Peer Discussion and Justification of Word Categorization

To increase accuracy and reliability, students participated in structured peer discussions to refine their word classifications. Peer discussions were structured as follows:

- Step 1: Students reviewed each other's word lists and provided a rationale for word classifications.
- Step 2: Disagreements were resolved through discussion, with students using examples, dictionary definitions, and corpus concordance lines to justify their choices.
- Step 3: Students documented their final classifications, along with written justifications, which were submitted as student notes and used as qualitative data in this study.

To ensure reproducibility, students were encouraged to cross-reference their findings with:

- Dictionaries (Oxford Online, Longman, Wikipedia for discipline-specific meanings)
- Corpus concordances (COCA, BAWE, SketchEngine word sketches)
- Discipline-specific textbooks and research articles

The peer discussion process was not audio-recorded due to ethical considerations but was systematically documented through student reflections and written justifications.

4.4. Data Analysis Procedures

The study employed a mixed-methods approach, integrating quantitative corpus analysis with qualitative peer discussion and student presentation insights.

4.4.1. Quantitative Analysis

- The Academic Word List (AWL)[6] was used as a reference to compare student-generated word lists.
- R programming was used to analyze word frequency and distribution, identifying the most pervasive AWL words in student corpora.
- A statistical comparison was conducted to examine variations in sub-technical vs. technical word usage across disciplines (see **Table 2**).

Table 2. CEW corpus compiled by Engineering students.

Category	Number of Words	Percentage	Corpus Size	Examples
Blockchain	356	95.7	0.8055M	area, range, complex, overall, dynamic, mechanism
Data Science	336	90.32	0.4553M	focus, positive, accurate, external, mode
Ethics and Privacy in Big Data	329	88.44	0.4560M	cycle, contact, scheme, enhance, transport
Film Art	329	88.44	0.7578M	media, brief, insight, chemical, rely
Blockchain	324	87.1	0.4890M	discrete, priority, exceed, profession, emphasis
Interest Rate Risk Management	320	86.02	0.4612M	aggregate, qualitative, document, trigger, complement
Software Defined Networks	316	84.95	0.8137M	react, hypothesis, rigid, proceed, fund
Machine Learning	299	80.38	0.2995M	factor, process, section, volume, energy
Robotics	295	79.3	0.3164M	normal, role, appropriate, impact, volume
Business Technology	293	78.76	0.2586M	issue, technical, input, perspective, strategy
Augmented Reality	290	77.96	0.3464M	component, generate, domain, medium, output
Computer Network	288	77.42	0.4833M	constant, benefit, crucial, abstract, maintain
Computer Vision	272	73.12	0.2845M	construct, prior, task, infrastructure, code
Networks	263	70.7	0.4066M	sequence, virtual, distinct, identical, file
Computer Networking	263	70.7	0.3621M	precise, valid, duration, principal, aware
Stereochemistry of organic compounds	253	68.01	0.4008M	denote, formula, proportion, reverse, widespread
Electric Vehicle	247	66.4	0.2457M	function, project, significance, technology, approach
Code Compilers	239	64.25	0.3878M	derive, logic, plus, adapt, attribute
Civil Engineering Materials	218	58.6	0.1841M	contribute, detect, brief, insight, finite
Additive Manufacturing	212	56.99	0.2658M	retain, induce, collapse, schedule, undergo
Biochemical Engineering	210	56.45	0.2192M	resolve, adjacent, behalf, preliminary, manual
Economics NCERT	206	55.38	0.1932M	constant, final, major, previous, specific
Physics and Astronomy	188	50.54	0.1516M	network, version, survey, fundamental, sum
Sodium Ion Battery	149	40.05	0.1548M	data, method, project, transfer, comprehensive
Cardiovascular Biology	136	36.56	0.0081M	similar, design, specific, theory, achieve
Aerodynamics and Wind Engineering	131	35.22	0.0707M	domain, feature, network, parameter, error
Geotechnical Engineering	83	22.31	0.0653M	esign, obtain, index, dynamic, institute
Semiconductors	70	18.82	0.0482M	transfer, image, parameter, principle, relevant
computational_chemistry	5	1.34	0.0221M	layer, compute, denote, implicit, predict

4.4.2. Qualitative Analysis

- Student notes and written justifications from peer discussions and presentations were coded manually using a content analysis framework.
- Coding focused on four major themes:
    1. Issues and dilemmas in categorizing words
    2. Sources and strategies used for classification
    3. Justifications for technical vs. sub-technical categorization
    4. Observations on disciplinary lexical patterns
- Representative excerpts were analyzed to illustrate common challenges and strategies in vocabulary classification (see Excerpts in the Results section).

4.5. Ethical Considerations

This study adhered to ethical research principles in classroom-based action research. While formal written consent was not obtained, student participation was voluntary, and their engagement in corpus compilation and vocabulary classification was part of regular coursework, ensuring no undue pressure to participate.

To protect participant anonymity and confidentiality:

- No personal identifiers were recorded in the corpus data or discussion transcripts.
- Data was anonymized and reported in aggregate form.
- Secure storage measures were implemented, restricting data access to the researcher.

Following best practices in educational action research[41, 42], this study involved minimal risk, as no sensitive personal data was collected.

5. Results

5.1. Pervasiveness of AWL Words

To answer the first research question on the pervasiveness of AWL in the CEW corpus, we used two tools: the AWL by Coxhead[6] and the *R* program. While the AWL lists were used as a reference to compare the raw lists generated by *SketchEngine* tools, the R Program was used to sort the words by their frequency and compare them with the AWL lists. **Table 3** below shows the number of words listed in different sub-lists of the AWL word list in the corpora built by the students. *R* programming sorted the tokens by their frequency of occurrence across all 29-word lists and filtered out the AWL words. The selected words should have occurred at least five times in at least ten-word lists. Finally, a total of 372 headwords out of 570 from the 10 sub-lists of the AWL were identified.

Table 3. Number of AWL words in each of the 29 sub-corpora in the CEW corpus.

Category	Number of Words	Percentage	Corpus Size	Examples
Blockchain	356	95.7	0.8055M	area, range, complex, overall, dynamic, mechanism
Data Science	336	90.32	0.4553M	focus, positive, accurate, external, mode
Ethics and Privacy in Big Data	329	88.44	0.4560M	cycle, contact, scheme, enhance, transport
Film Art	329	88.44	0.7578M	media, brief, insight, chemical, rely
Blockchain	324	87.1	0.4890M	discrete, priority, exceed, profession, emphasis
Interest Rate Risk Management	320	86.02	0.4612M	aggregate, qualitative, document, trigger, complement
Software Defined Networks	316	84.95	0.8137M	react, hypothesis, rigid, proceed, fund
Machine Learning	299	80.38	0.2995M	factor, process, section, volume, energy
Robotics	295	79.3	0.3164M	normal, role, appropriate, impact, volume
Business Technology	293	78.76	0.2586M	issue, technical, input, perspective, strategy
Augmented Reality	290	77.96	0.3464M	component, generate, domain, medium, output
Computer Network	288	77.42	0.4833M	constant, benefit, crucial, abstract, maintain
Computer Vision	272	73.12	0.2845M	construct, prior, task, infrastructure, code
Networks	263	70.7	0.4066M	sequence, virtual, distinct, identical, file
Computer Networking	263	70.7	0.3621M	precise, valid, duration, principal, aware
Stereochemistry of organic compounds	253	68.01	0.4008M	denote, formula, proportion, reverse, widespread
Electric Vehicle	247	66.4	0.2457M	function, project, significance, technology, approach
Code Compilers	239	64.25	0.3878M	derive, logic, plus, adapt, attribute
Civil Engineering Materials	218	58.6	0.1841M	contribute, detect, brief, insight, finite
Additive Manufacturing	212	56.99	0.2658M	retain, induce, collapse, schedule, undergo
Biochemical Engineering	210	56.45	0.2192M	resolve, adjacent, behalf, preliminary, manual
Economics NCERT	206	55.38	0.1932M	constant, final, major, previous, specific
Physics and Astronomy	188	50.54	0.1516M	network, version, survey, fundamental, sum
Sodium Ion Battery	149	40.05	0.1548M	data, method, project, transfer, comprehensive
Cardiovascular Biology	136	36.56	0.0081M	similar, design, specific, theory, achieve
Aerodynamics and Wind Engineering	131	35.22	0.0707M	domain, feature, network, parameter, error
Geotechnical Engineering	83	22.31	0.0653M	esign, obtain, index, dynamic, institute
Semiconductors	70	18.82	0.0482M	transfer, image, parameter, principle, relevant
computational_chemistry	5	1.34	0.0221M	layer, compute, denote, implicit, predict

As can be seen, some corpora such as the Blockchain and the Data Science contained a relatively large number of AWL words. In fact, most of the texts from the sub-domains of the computer science discipline contained a fairly large number of AWL words while the corpus of computational chemistry and semiconductors contained a relatively smaller number of AWL words. The discipline of computational chemistry is densely organized around formulas and other forms of numerical data while the corpora on CS includes a higher proportion of argumentation. **Table 4** presents the most frequently occurring words from the Academic Word List (AWL) across the student-compiled sub-corpora.

Table 5 shows the sub-technical vocabulary identified by the students in three different disciplines/themes. While many words listed under Aerodynamics and Interest-rate risk management using statistical methods are very technical (TAM 4 or Step 3) in nature as their specialized senses are least related to their general senses and outside the purview of disciplines these meanings occur minimal; however, the words listed under the Big Data domain seem pervasive and general due to the popularity of discourses of computer science (CS) discipline have gained in other registers such as newspapers and magazines. Moreover, the specialized meanings projected by words such as *privacy, data, protection,* and *security* are in fact, the most frequent and general senses of the words. Based on the observation that CS domains use a large number of general and AWL words (see **Table 3**), it may be reasonable to argue that the density of very technical vocabulary in CS is relatively lesser. It is, however, important to discuss why the students have listed some of the AWL words under the sub-technical category.

Table 4. Distribution of the most frequently used AWL words in the 29 sub-corpora.

Rank	*Words from AWL*	*Distribution in the Sub-Corpora*
1	area	28
2	available	28
3	data	28
4	factor	28
5	function	28
6	method	28
7	physical	28
8	potential	28
9	process	28
10	project	28
11	range	28
12	region	28
13	research	28
14	section	28
15	significant	28
16	similar	28
17	structure	28
18	technology	28
19	volume	28
20	approach	27

Some students used certain strategies in identifying the sub-technical words. For example, words listed under the topic Big Data were identified based on their occurrence outside CS: According to this student, "the text would lose critical pieces of information without these words while they may also appear in other domains in general." Also, words expressing a specialized sense in one discipline and having a general meaning were also categorized under the sub-technical. A student who compiled a corpus of additive manufacturing says:

(1) … for example, the word *feedstock* is used in many disciplines and its meaning differs from field to field but in our area of interest, it means the amount of material available at any specific point of time during the process of manufacturing. (Student Notes)

Another student mentioned that one of the conditions for selecting a word under the technical category was its potential unuse in other disciplines. For example, the words listed in **Table 6** show the technical words in their respective corpora. Interestingly, many of the technical words were nouns projecting discipline specific meanings. However, it is interesting to note how words like *cloud, reidentify,* and *privacy-related* were listed under the technical category while words like *electrode, profitability,* and *liquidity* were listed under the sub-technical category. One strategy consistently used to identify the technical words was to verify the consistency of the literal meaning across the disciplines.

(2) For example, with words like *throughput, checksum*, and *bandwidth*, I felt that they fell more in the sub-technical territory, whereas Hrishabh felt that they were more technical since their meaning stayed the same across all domains. He managed to get his point across and we ultimately classified them as Technical. (Student Presentation)

Table 5. Sub-technical words in three different sub-corpora of the CEW corpus.

Ethics and Privacy in Big Data	**Aerodynamics**	**Interest-Rate Risk Management Using Statistical Methods**
data	additive	hedging
information	oscillate	volatility
research	exchanger	bond
technology	porous	reserve
risk	velocity	banking
big	phase	funds
issue	pouch	duration
privacy	dispersibility	deposits
security	thickness	capital
ethic	reciprocation	liquidity
city	solicitation	risks
study	detonation	regression
health	kinematic	liabilities
model	longitudinal	correlation
system	gradient	coefficient
project	lifespan	swaps
approach	simulative	maturities
protection	electrode	investors
concern	parasitic	profitability
science	ventilation	differential

Table 6. Technical words in three sub-corpora of the CEW corpus.

Chemistry: Sodium-Ion Battery	Biology	Ethics and Privacy in Big Data
acetylene	artery	Privacy-preserving
sodium-ion	doppler	Anonymization
electrochemical	stenosis	Homomorphic
anode	hypertension	Autoencoder
nanosheet	ultrasound	K-anonymity
cathode	angina	Reidentification
redox	angiography	L-diversity
electrolyte	diabetes	Denoising
lithium-ion	coronary	Ciphertext
anionic	echocardiography	Cloud
galvanostatic	tomography	Reidentify
sodium	hypoxia	Homomorphic
electrode	cardiology	Representativity
lithium	infarction	Paralinguistics
bismuth	arteriolar	Privacy-enhanced
coulombic	vasculature	Interpretability
calcination	ischemia	Privacy-related
conductivity	cardiomyopathy	Nosy neighbour (type of attack mechanism)

In determining the technicalness of the individual words, the students had to discuss their observations and analysis with their peers. In order to reduce the burden of analysing a large number of words, we asked the students to identify at least 50 words for the sub-technical category and 50 words for the technical category and justify their reasons. In what follows, we analyse the data obtained from the student notes and oral presentations to understand better the ways the students categorized the words.

5.2. Development of Sub-Technical and Technical Word Lists

To address the second question on how the students utilized the technicalness scale and determined the technicality of the words, we analysed the student notes, which included student deliberations on the ways they processed the word lists. Most of the notes were written in the form of explanatory introductions broadly containing the following moves [43]:

a) introduce the corpus,
b) present an issue(s) in sorting the words,
c) present the strategies/sources used to address the issue or explain the procedures of addressing the issue, and
d) provide a commentary/an evaluation.

All the student notes and the transcriptions of presentations were coded manually for the moves used. The codes were decided *a posteriori* based on the functions they served [44, 45]. Majorly, four codes emerged: *issues, strategies, procedures,* and *comments*.

a Issues and dilemmas in categorizing the words
b Sources used in categorizing the words (what & how)
c Strategies used in categorizing the words (what & how)
d Comments on the developed word lists

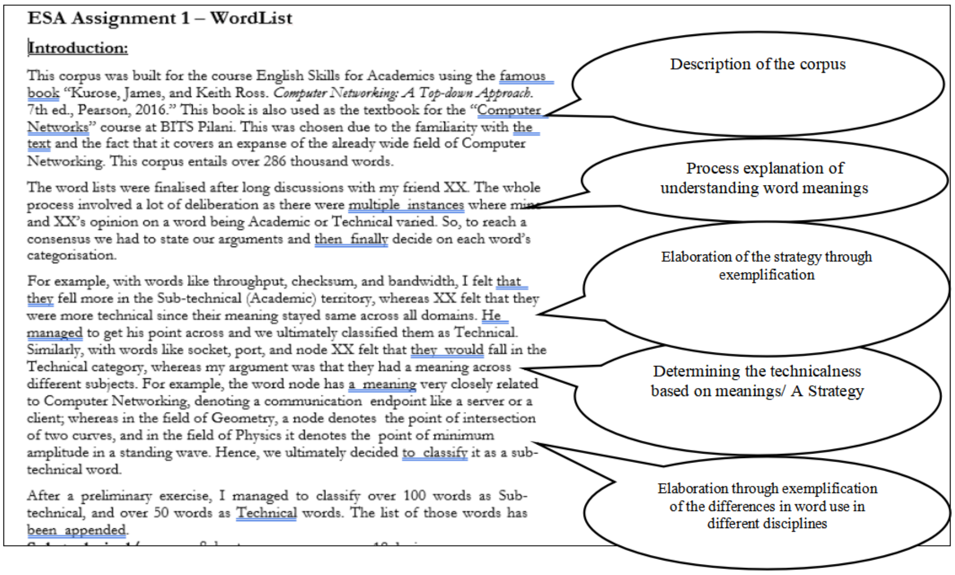
Figure 3 and **Figure 4** demonstrate the process of coding the student notes and the presentations.

a. Issues and Dilemmas in Categorizing the Words

One of the issues that ignited a discussion among the peers is the interdisciplinary lexical borrowing, which is common when two or more disciplines engage in a disciplinary dialogue (biochemistry: biology-chemistry; robotics: computer science and mechanical engineering). This has caused a discussion among students, particularly when a word being used frequently in other allied disciplines did not necessarily project a different specialized meaning. This led to the dilemma whether such words with high frequency across disciplines projecting a singular meaning to be categorized under 'sub-technical' or 'technical'. Some students, however, marked them under the sub-technical category while some categorized them under the technical. The following explains the reasons why such words were classified under the sub-technical.

(3) … issue that we had to face was to decide amongst the words which were technical as per the field of Machine

Learning but were also used in Probability Theory and other sub-branches of Mathematics. This was due to the fact that ML works very closely in relation to probability theory and draws most of its concepts from linear algebra. Thus a lot of seemingly technical words like 'k-means' 'linear regression' and 'meta-analysis' were recategorized into the Sub-technical Words. (Source: Student Notes)


ESA Assignment 1 – WordList

Introduction:

This corpus was built for the course English Skills for Academics using the famous book "Kurose, James, and Keith Ross. *Computer Networking: A Top-down Approach*. 7th ed., Pearson, 2016." This book is also used as the textbook for the "Computer Networks" course at BITS Pilani. This was chosen due to the familiarity with the text and the fact that it covers an expanse of the already wide field of Computer Networking. This corpus entails over 286 thousand words.

The word lists were finalised after long discussions with my friend XX. The whole process involved a lot of deliberation as there were multiple instances where mine and XX's opinion on a word being Academic or Technical varied. So, to reach a consensus we had to state our arguments and then finally decide on each word's categorisation.

For example, with words like throughput, checksum, and bandwidth, I felt that they fell more in the Sub-technical (Academic) territory, whereas XX felt that they were more technical since their meaning stayed same across all domains. He managed to get his point across and we ultimately classified them as Technical. Similarly, with words like socket, port, and node XX felt that they would fall in the Technical category, whereas my argument was that they had a meaning across different subjects. For example, the word node has a meaning very closely related to Computer Networking, denoting a communication endpoint like a server or a client; whereas in the field of Geometry, a node denotes the point of intersection of two curves, and in the field of Physics it denotes the point of minimum amplitude in a standing wave. Hence, we ultimately decided to classify it as a sub-technical word.

After a preliminary exercise, I managed to classify over 100 words as Sub-technical, and over 50 words as Technical words. The list of those words has been appended.

Figure 3. Example of move-coding on student notes.

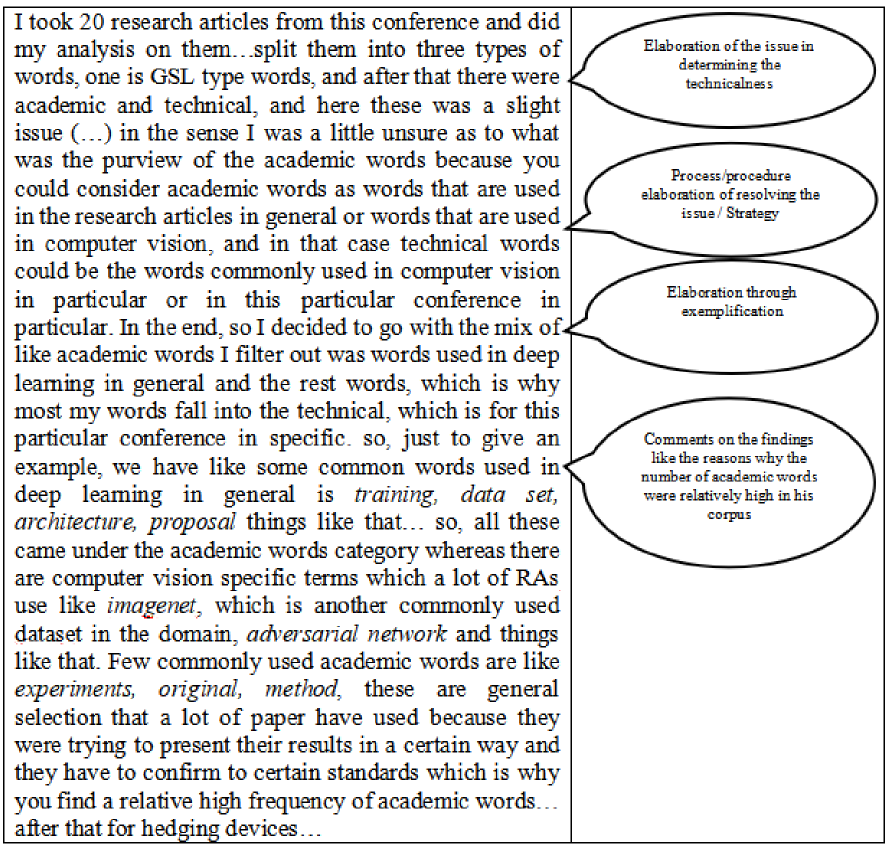

I took 20 research articles from this conference and did my analysis on them…split them into three types of words, one is GSL type words, and after that there were academic and technical, and here these was a slight issue (…) in the sense I was a little unsure as to what was the purview of the academic words because you could consider academic words as words that are used in the research articles in general or words that are used in computer vision, and in that case technical words could be the words commonly used in computer vision in particular or in this particular conference in particular. In the end, so I decided to go with the mix of like academic words I filter out was words used in deep learning in general and the rest words, which is why most my words fall into the technical, which is for this particular conference in specific. so, just to give an example, we have like some common words used in deep learning in general is *training, data set, architecture, proposal* things like that… so, all these came under the academic words category whereas there are computer vision specific terms which a lot of RAs use like *imagenet*, which is another commonly used dataset in the domain, *adversarial network* and things like that. Few commonly used academic words are like *experiments, original, method,* these are general selection that a lot of paper have used because they were trying to present their results in a certain way and they have to confirm to certain standards which is why you find a relative high frequency of academic words… after that for hedging devices…

Figure 4. Example of move-coding on student presentations.

Similarly, some students predicted the potential use of a word outside their corpus and explained why some words were in the grey area. For example, a pharmacy major student, who was studying courses in biology and chemistry disciplines, analysed a corpus of biochemistry research articles, and explained her preference for listing some words under the sub-technical category in the following way.

(4) Certain words encountered in this list were in a grey area between the two categories. The word 'zeolite', for example, refers to a material that is noted for its ability to facilitate the exchange of 'ions' and this means that it can potentially be used in fields other than biology, in cases such as materials engineering and chemistry. (Source: Student Presentation)

However, a computer science student tried to rationalize the technicalness scale using her disciplinary knowledge in the following way.

(5) Technical words like *cloud* and *nosy neighbour* have different meanings in different contexts but take a more technical meaning in this domain. *Anonymity* is classified as an academic word but *Anonymization* is a technical term since it describes the data process technique.

For many students, their intuition and prior exposure to disciplinary texts played a role in deciding the technicalness of the word. When they sensed that a word projected a specialized sense in their discipline, they tried to categorize it under the technical words. However, they have thoroughly verified their preferences by consulting various sources.

b. Sources and Strategies Used in Categorizing the Words

Three online reference sources were majorly consulted: online dictionaries, Wikipedia, concordances from the Corpus of Contemporary American English (COCA) and the sub-corpora of CEW.

(6) In case of a conflict, the resolution happened either by giving examples and counterexamples or by using tools such as COCA/searching the internet for the meaning and academic use of the terms. For example, the words like *sodium, carbon, ion* should be classified as Technical according to me. Still, in reality, they are classified as semi-technical words because they are frequently used, and many people can comprehend their use and meaning. (Student Notes)

One pair of students majoring in Biology redefined their criteria after deliberations. On more than one occasion, they had consulted the Oxford Online Dictionary to specifically find instances of the word being used in other disciplines. When they found words projecting specialized meanings in two different disciplines such as *organ, membrane, muscle*, and *capillary,* they categorized them under the sub-technical vocabulary because they were 'also found in texts other than biology.' Because the words *dobutamine, revascularization, plasminogen* were not found in other disciplines, they identified them as technical words. Similarly, a student who was examining the word *equilibrium* in a dictionary in two different contexts said the following:

(7) This led to quite interesting discussions between my friend and myself concerning classifying words to sub-technical: the word *equilibrium* is used in economics when the supply and the demand are equal, but the word equilibrium in physics is defined as a state in which opposing forces or influences are balanced. So this is a sub-technical word. (Student Presentation)

While intuition and prior exposure to vocabulary in other disciplinary contexts played a role in determining the technicalness of the words, accessing resources such as online dictionaries, blogs, and Wikipedia strengthened their assumptions. Some students mentioned that they consulted dictionary definitions and examples for both contextual usage and polysemy. Some students used the strategies of searching for meanings beyond their disciplines and examining the collocations of the search word.

(8) Evaluation of the word categories was done (…) cross-referencing from Google's English Dictionary provided by Oxford Languages. From the discussion with Yash, I reclassified words like Bootstrap (1. Lit. get into or out of a circumstance using available resources, 2. A concept in web development), Interpolate, Differential (Lit. Difference between amount of things), Eigenvalue (Concept used across Physics, Mathematics, and Computer Science), and many more… (Student Notes)

(9) Conflicts arising are addressed by looking through the concordances of the word, exploring the internet and putting forwards suitable examples for the same. (Student Notes)

Among other corpus tools, students frequently used

SketchEngine's word sketch to examine the collocates and *n*-gram tools to study the patterns. Some of their explanations were based on the patterns the words formed. For example, a student mentioned that she and her peer looked up the word *area* from the sublist 1 of the AWL in her corpus for the most frequent phraseological patterns. They discovered that the word area had occurred frequently in a rather transparent construction 'in the *area* of + noun' (13 hits) in Block Chain Technology while it occurred only in the construction 'the effective detection *area*' (5 hits) in the Astronomy corpus. By comparing the word sketches generated by the *SketchEngine* tool they determined the technicalness of certain words. Similarly, another student examined the word *parallel* in two different corpora and determined its technicality based on its collocates. **Figure 5**, taken from a student's notes, is given below.

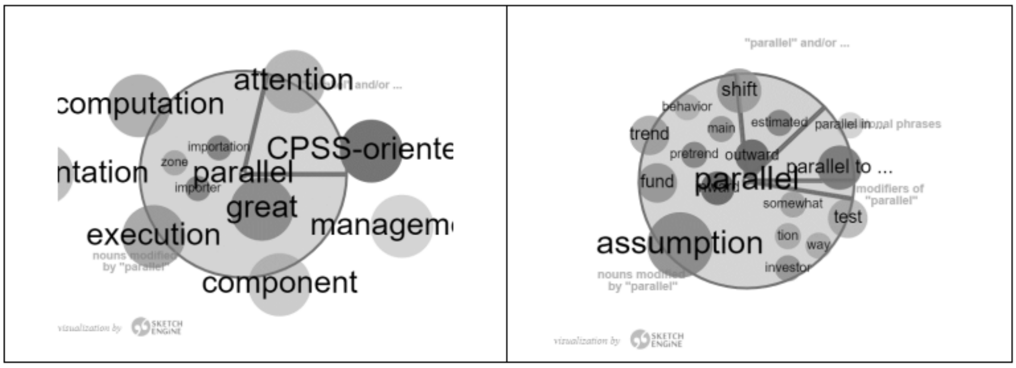


Figure 5. Word sketch output for the word 'parallel' in two different sub-corpora of SCWE generated by a student.

In some cases, lexical relations shown by word sketch analysis were discussed with reference to their prior exposure to disciplines. The following excerpt from the transcription of the presentations showed how the lexical relations were identified in terms of their collocations and thematic relevance.

(10) … tools that help in the process of *casting*, like *spruce, chills, riser, mould* etc., and the ends of different processes that happen during the process of casting like *cooling, solidification, melting,* and some general scientific words like *temperature, pressure, mass, volume* were heavily used in the research papers and in the casting industry. And finally the types of casting like there are various types of casting, but these three types of casting were the most frequent in the research papers and are also heavily used in the casting industry, like *die-casting, continuous casting,* and *squeeze casting*. (italicized by the authors)

(Student Presentation)

c. Strategies Used in Categorizing the Words

One of the most frequently used strategies was to draw on their knowledge of disciplinary courses they were exposed to justify the technicalness of a word. The focus of peer discussions, as reported in the notes, was usually about citing the meanings they were introduced to in other courses. Some of them examined multiple instances of word use for the probability of the word being dependent on other words for its meanings. In some cases, if the word's meaning remained consistently the same across the disciplines/ examples, they considered it highly technical even though the word appeared outside their discipline. However, for students who have analysed a corpus with CS and Management related domains, many words are either academic or sub-technical, as in the case (11), because they have become part of 'normal person's life.' For example, if we go by the stepwise classification of Chung and Nation, the word *linear regression* would be a technical word. However, since this term is often used in reporting research across disciplines, students categorized it under the sub-technical. Similarly, the mineral *zeolite, bioinformatics,* and *hyperplane* were categorized under the sub-technical vocabulary category.

(11) The classification of *routing*, as well as *router*, are two terminologies that have sparked considerable debate between us. The term 'route' is a broad phrase with a meaning that extends beyond Networks. However,

the name ’router,’ which is derived from this word, is unique to this sector. Because the word router has become such an integral part of the normal person’s life, it may be used without further elaboration. As a result, I chose to categorize router as an academic term. (Student Notes)

When confronted by a set of words that could cause confusion, students tried to clearly articulate their categorization boundaries. One such instance is given below.

(12) Segregation into sub-technical was made on the following principles:

- Words must be recognizable as entities from the particular choice of discipline or have their origins rooted in that discipline
- The word must provide no exclusive context of a specific discipline when used in any statement of other subject areas.
- The word could be minimally identified in general contexts

Segregation into technical was done on the following basis:

- Words must be recognizable as entities of a particular subject domain
- The words must additionally be of very little to no significance in contexts apart from the particular discipline. In other words, any use of such words in linguistic sentences must redirect its underlying morphology to a specific academic field

(Student Notes)

d. Comments on the Developed Word Lists

Student comments or findings about the technical words were interesting. Comments included generalizations, observations about words, and reasons why they categorized certain words under a category. Some students tried to group the technical words into categories such as the processes, names of some of the algorithms that are normally used for image classification. A student who was working on a project titled ‘Image classification of remotely sensed data’ explained why he had categorized some of the otherwise academic words under the technical category.

(13) …five most used words in Image Classification are *classification, remote, sensing, training* and *class*. Even though they are very commonly used words, they gain great significance in terms of technicality. So even though there were many more technical words like *semi-supervised* and *clustering*, still, these happened to be the most used technical words. (Student Presentation)

A comment showing the distinction between the sub-technical and technical words is interesting. According to this student, sub-technical words are used to ‘make the readers comfortable’ while technical words are used mainly in research papers “whose main concern is to provide the insights of the respective research done.”

While disagreements over the technicalness of a word continue to allure students to categorize many words under the semi-technical category, students seemed to believe in the view that if a word tended to show some degree of context specificity in terms of its meaning, they preferred to identify them as technical words.

6. Discussion

In this study, we aimed to explore how undergraduate students in ESL science and engineering programs assess the technicality of words using a technicalness scale, with a focus on the development of the scale and the role of peer discussions in vocabulary development. Our findings indicate that personalized learning—where students engage with vocabulary relevant to their disciplines—can enhance their awareness of disciplinary-specific terminology and encourage deeper exploration of words in context.

The first research question examined the prevalence of AWL words in a corpus of student-generated texts. Our analysis revealed an uneven distribution of AWL words across disciplines, challenging the notion of a universal lexical repertoire, such as the AWL [2]. Certain fields, like semiconductors and computational chemistry, were found to be highly technical, while other domains, such as Blockchain and Data Science, featured a higher frequency of AWL words. This observation suggests that relying solely on general academic vocabulary as a foundation for a universal academic lexicon may not be appropriate. Instead, we advocate for a broader conception of lexical repertoires that includes not only general academic words but also sub-technical and technical vocabulary [27].

Our second research question focused on how students processed and categorized vocabulary using the technicalness scale. Categorizing words based on their technicality proved to be a complex task, as students employed various strategies to determine whether a word fit into the academic, sub-technical, or technical categories. These strategies included examining word context, cross-referencing dictionary entries, and comparing occurrences across corpora. Word form also emerged as a key factor in determining technicality, with nouns being more commonly classified as technical words. Furthermore, disciplinary overlap and the increasingly multi-disciplinary nature of research contributed to students' perceptions of technicality, as some words that appeared discipline-specific were also used outside their respective fields.

The findings of this study underscore the value of peer discussions in vocabulary development. Through peer review and collaborative analysis, students brought their disciplinary knowledge into the process of classifying words, enriching the form-meaning relationships they identified. This process not only facilitated their understanding of technical and academic vocabulary but also contributed to a more personalized learning experience, as students related the vocabulary they encountered to their own academic contexts.

The implications for English for Academic Purposes (EAP) teaching are significant. Our study suggests that integrating dictionary-based exercises, discipline-specific writing tasks, and peer review into the curriculum can help students develop a more specialized and restricted lexical repertoire, focusing on both general academic words and technical vocabulary. Moreover, we recommend further research into how students relate to and acquire technical vocabulary within their disciplines, with an emphasis on incorporating academic literacy models that draw on the epistemologies of disciplinary discourses (Lea, 2004). The role of peer review in vocabulary study, as evidenced in this study, appears to be crucial in enhancing students' understanding of form-meaning relationships, while simultaneously linking the vocabulary they study to their academic and disciplinary experiences.

7. Conclusions

A single lexical repertoire for all the students is only possible when the students themselves take the initiative to determine their lexical needs. As EAP teachers, we should facilitate the process by providing them with necessary tools (e.g., corpus tools, the technicalness scale) and skills for using and interpreting them. This study has tried to do the same and capture the process of analysing the raw word lists. By handing over the baton of responsibility to the students and by showing them the way they can use the corpus in their own learning, we have integrated the DIY corpus research (Do-It-Yourself) into our EAP course [46–48].

Author Contributions

Conceptualization, S.T., V.C.; methodology, R.Y., V.C.; software, R.Y.; formal analysis, R.Y., V.C.; investigation, R.Y., S.T., V.C.; resources, S.T., V.C.; data curation, R.Y..; writing—original draft preparation, R.Y., V.C.; writing—review and editing, R.Y., S.T., V.C.; visualization, V.C.; supervision, S.T., V.C.; project administration, V.C. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Informed Consent Statement

Not applicable.

Institutional Review Board Statement

This study did not require Institutional Review Board (IRB) approval as it involved minimal-risk, classroom-based research within the standard educational framework. No personal identifiers were collected, and participation was voluntary, ensuring ethical compliance with established research guidelines.

Data Availability Statement

The dataset, consisting of compiled word lists from student-generated corpora, is available in digital format and can be shared upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Gardner, D., Davies, M., 2014. A new academic vocabulary list. Applied Linguistics. 35(3), 305-327.
[2] Hyland, K., Tse, P., 2007. Is there an "academic vocabulary"?. Teaching English to Speakers of Other Languages Quarterly. 41(2), 235-253.
[3] Nation, I.S.P., 2001. Learning Vocabulary in Another Language. Cambridge University Press: Cambridge, UK. pp. 13-21, 217-235.
[4] Nation, I.S.P., 2016. Making and Using Word Lists for Language Learning and Testing. John Benjamins Publishing Company: Amsterdam, Netherlands. p. 116.
[5] Bi, J., 2020. How large a vocabulary do Chinese computer science undergraduates need to read English-medium specialist textbooks?. English for Specific Purposes. 58, 77-89.
[6] Coxhead, A., 2000. A new academic word list. Teaching English to Speakers of Other Languages Quarterly. 34(2), 213-238.
[7] Coxhead, A., Demecheleer, M., 2018. Investigating the technical vocabulary of plumbing. English for Specific Purposes. 51, 84-97.
[8] Yang, M.N., 2015. A nursing academic word list. English for Specific Purposes. 37, 27-38.
[9] Li, Y., Qian, D.D., 2010. Profiling the Academic Word List (AWL) in a financial corpus. System. 38(3), 402-411.
[10] Masrai, A., Milton, J., 2018. Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. Journal of English for Academic Purposes. 31, 44-57.
[11] Charles, M., 2012. 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. English for Specific Purposes. 31(2), 93-102.
[12] Ackermann, K., Chen, Y.H., 2013. Developing the Academic Collocation List (ACL) - A corpus-driven and expert-judged approach. Journal of English for Academic Purposes. 12(4), 235-247.
[13] Valipouri, L., Nassaji, H., 2013. A corpus-based study of academic vocabulary in chemistry research articles. Journal of English for Academic Purposes. 12(4), 248-263.
[14] Biber, D., 2006. University Language: A Corpus-Based Study of Spoken and Written Registers. John Benjamins Publishing Company: Amsterdam, Netherlands. pp. 1-50.
[15] Biber, D., Gray, B., 2010. Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. Journal of English for Academic Purposes. 9(1), 2-20.
[16] Durrant, P., 2016. To what extent is the Academic Vocabulary List relevant to university student writing?. English for Specific Purposes. 43, 49-61.
[17] Jaidev, R., Chan, P., 2018. Embedding communication in the disciplines: A tale of two faculties. Innovation in Language Learning and Teaching. 12(3), 199-211.
[18] El-Dakhs, D.A.S., 2018. Why are abstracts in PhD theses and research articles different? A genre-specific perspective. Journal of English for Academic Purposes. 36, 48-60.
[19] Yin, S., Gao, Y., Lu, X., 2021. Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. System. 97, 102427.
[20] Bancroft-Billings, S., 2020. Identifying spoken technical legal vocabulary in a law school classroom. English for Specific Purposes. 60, 9-25.
[21] Muñoz, V.L., 2015. The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. English for Specific Purposes. 39, 26-44.
[22] Anthony, L., 2020. AntConc. Available from: www.laurenceanthony.net/software/antconc/
[23] Kim, C., Crosthwaite, P., 2019. Disciplinary differences in the use of evaluative that: Expression of stance via that-clauses in business and medicine. Journal of English for Academic Purposes. 41, 100775.
[24] Dang, T.N.Y., Coxhead, A., Webb, S., 2017. The academic spoken word list. Language Learning. 67(4), 959-997.
[25] Lei, L., Liu, D., 2016. A new medical academic word list: A corpus-based study with enhanced methodology. Journal of English for Academic Purposes. 22, 42-53.
[26] Drayton, J., Coxhead, A., 2023. The development, evaluation and application of an aviation radiotelephony specialised technical vocabulary list. English for Specific Purposes. 69, 51-66.
[27] Ha, A.Y.H., Hyland, K., 2017. What is technicality? A technicality analysis model for EAP vocabulary. Journal of English for Academic Purposes. 28, 35-49.
[28] Chung, T.M., Nation, P., 2003. Technical vocabulary in specialised texts. Reading in a Foreign Language. 15(2), 103-116.
[29] Chung, T.M., Nation, P., 2004. Identifying technical vocabulary. System. 32(2), 251-263.
[30] McLaughlin, E., Parkinson, J., 2018. 'We learn as we go': How acquisition of a technical vocabulary is supported during vocational training. English for Specific Purposes. 50, 14-27.
[31] Kay, A.E., Hardy, J., Galloway, R.K., 2018. Learning from peer feedback on student-generated multiple-choice questions: Views of introductory physics students. Phys Rev Phys Educ Res. 14(1), 010119.

DOI: https://doi.org/10.1103/PhysRevPhysEducRes.14.010119
[32] Filius, R.M., de Kleijn, R.A.M., Uijl, S.G., et al., 2018. Strengthening dialogic peer feedback aiming for deep learning in SPOCs. Computers & Education. 125, 86-100.
[33] Yang, M., Badger, R., Yu, Z., 2006. A comparative study of peer and teacher feedback in a Chinese EFL writing class. Journal of Second Language Writing. 15(3), 179-200.
[34] Ertmer, P.A., Richardson, J.C., Belland, B., et al., 2007. Using peer feedback to enhance the quality of student online postings: An exploratory study. Journal of Computer-Mediated Communication. 12(2), 412-433.
[35] Topping, K., 1998. Peer assessment between students in colleges and universities. Review of Educational Research. 68(3), 249-276.
[36] Lubis, A.H., 2020. Corpus-driven genre pedagogy of English for research publication purposes: Toward an ELF-referenced critical teaching framework. International Journal of Education. 12(2), 109-121.
[37] Hussain, Z.D., 2021. Incorporating digital corpus analysis technology into English for Academic Purposes (EAP) lesson design for the Sri Lankan tertiary context: A self-experimentation by a ‘novice’. Advances in Technology. 1(3), 483-507.
[38] Ferdiansyah, S., Picard, M., 2024. Modeling expression of the research gap in text-based reading instruction: A corpus-based approach. Journal of Second Language Studies. 7(2), 295-319.
[39] Nesi, H., 2016. Corpus studies in EAP. In: Hyland, K., Shaw, P. (eds.). The Routledge Handbook of English for Academic Purposes. Routledge: London, UK. pp. 206-217.
[40] Lee, D., Swales, J., 2006. A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. English for Specific Purposes. 25(1), 56-75.
[41] Burns, A., 2009. Doing Action Research in English Language Teaching: A Guide for Practitioners, 1st ed. Routledge: New York, USA. pp. 70-75. DOI: https://doi.org/10.4324/9780203863466
[42] Mckay S.L., 2006. Researching Second Language Classrooms, 1st ed. Routledge: New York, USA. pp. 45-50. DOI: https://doi.org/10.4324/9781410617378
[43] Gardner, S., 2012. Genres and registers of student report writing: An SFL perspective on texts and practices. Journal of English for Academic Purposes. 11(1), 52-63.
[44] Moreno, A.I., Swales, J.M., 2018. Strengthening move analysis methodology towards bridging the function-form gap. English for Specific Purposes. 50, 40-63.
[45] Charles, M., 2011. Adverbials of result: Phraseology and functions in the Problem-Solution pattern. Journal of English for Academic Purposes. 10(1), 47-60.
[46] Zhao, H., Zhan, H., 2020. The use of genitive citations in academic writing. Lingua. 234, 102769.
[47] Chintalapalli, V., 2020. ‘Do-it-yourself’ English courses for higher education. Journal of English Language Teaching. 62(2), 10-14.
[48] Charles, M., Hadley, G., 2022. Autonomous corpus use by graduate students: A long-term trend study (2009-2017). Journal of English for Academic Purposes. 56, 101095.